Artificial Intelligence in Malware - Cop or Culprit?

Pan Juin Yang Jonathan and Chun Che Fung

School of Information Technology, Murdoch University, Perth, WA

Email: {Jonathan.Pan.JY@gmail.com | l.fung@murdoch.edu.au}

Abstract—**Malware is very much a part of today's digital society as well as the battle against the malicious attacks. Victory over this struggle is essential to ensure the proper functioning and efficient operations of the world's digital economy. The use of artificial intelligence in this virtual battle is vital. Malware has been noted to have many intelligent features like the ability to deceive their targeted victims and stealth capabilities to prevent detection. Similarly, anti-malware solutions leverage on artificial intelligence techniques to identify new malware threats and to keep the existing pool of malware at bay. This survey paper highlights how artificial intelligence is being used in information security specifically in both malware and anti-malware warfare.**

I. Introduction

Malware is very much a part of today's digital society whether we like it or not. Similarly, it is an ongoing battle to defend and combat against them [1]. Malware is any software that contains code with malicious intentions to inject themselves into the computer systems with or without the owner's consent. Typically they are disguised in the form of spam emails that flood the email accounts. It is not unsurprising that malicious programs may reside in a large number of computers at work places and homes. Such malware is popularly known as viruses and worms that continually introduce inconvenience and disruption to the daily computer operation. The battle to eradicate these malwares has led to many technological development and communities from commercial to research entities. All are working constantly in developing various forms of anti-malware solutions and defence strategies. However there are occasions where new strands or form of malware have spread rampantly across the Internet rendering the anti-malware software or solutions useless [2]. The battle between the good and the evil (depends which side of the fence) is an ongoing tug-of-war with no obvious end in sight. On this front, artificial intelligence (AI) may provide some leverages against the malware. AI is based on theories of computer science to enable reasoning, knowledge acquisition, planning, learning, perception and the ability to manipulate objects and knowledge. Hence, one may ask, "Does smart malware carry some forms of intelligence with the intention to outwit and to defeat the barrier of information security defences?", or, "How can the good guys use intelligent technologies to fend off the assault of malware and protecting the interests of individuals and organisations?"

Authors' note: Jonathan Pan has been accepted for admission as a Doctor of Information Technology student at Murdoch University, WA. He is currently residing in Singapore and working in the information security industry. This is a position paper on his research proposal.

The objective of this survey paper is to study how artificial intelligence is used in this virtual battle between the bad (malware) and good (defenders against malware). This paper first covers the current state of the struggle against malware. The paper will then cover how artificial intelligence is being used in both camps. Finally, this paper will explore research opportunities to incorporate intelligence techniques into the virtual weapons of the cyber warfare.

II. Malware Epidemic

Malware typically exists in the form of email spams, viruses, trojan horses, software rootkits, browser hijackers and worms. The capabilities of current day malware started first as a research study by Fred Cohen [3] in order to study how programs can infect other programs. However, others took this development further to include malicious code with the intent to infect host PCs with immediate or deliberately delayed outcomes. This could exist as a time bomb or becoming part of a *botnet* in order to participate in future attacks on other targets. Malware is a powerful enabler to crime. Malware has now been used in a wide range of ways to enable criminals to achieve their intentions for financial gains and other illegitimate objectives.

Organizations face significant risks if malware is not managed adequately. For example, Estonia's national internet capabilities were crippled by the onslaught attacks of malware in May 2007 [4]. In fact, the world at large is at risks. In the recent OECD meeting, a report titled '*Malicious Software (Malware): A Security Threat to the Internet Economy*' [5] highlights the need to have a strategy for global partnership against malware as the latter poses 'a serious threat to the Internet economy and to national security in the coming years'. Managing malware or preventing them from achieving its objectives is now an on-going war that requires immediate attention.

Is the world experiencing a malware epidemic and when will the next malware induced disaster occur? Will there ever be an end to this phenomenon? To address these concerns, there are researchers and engineers working hard to fend off such malware attacks. They come from a wide range of industries from academia to law-enforcement agencies to commercial companies. Similarly, there are many developers of malware from amateurs seeking thrills and fame, to serious organized cyber criminals', according to the OECD report [5].

Ever since the first virus developed by Cohen in 1983 as part of a research study, many new technological advancement have been introduced into malware development. A key advancement is the inclusion of artificial intelligence into malware. This advancement also took place in the counter measures against malware [6]. Artificial intelligence techniques have been studied and introduced into detection and prevention mechanism in the malware war. The motivation behind such extensive research by the academia into malware advancement and likewise their counter measures is to facilitate the identification of threats that may occur in completely new paradigm as Fernandez and Bureau [23] explains before it may happen.

In order to study the roles of AI in this war, it is important to understand the characteristics of intelligent software and how they are being classified as "intelligent". There are well known AI technologies, methodologies and developed systems such as expert systems (ES), soft computing, neural networks (NN), genetic algorithms (GA), fuzzy systems, and computational intelligence (CI) techniques. Another way to assess the intelligence of software is whether the software mimics biological behaviours [7] like the ability to mutate, propagate, infect its host, detect detection, overcome counter measures in the digital world much like the real-world. Turing proposes a test of a machine's ability to demonstrate intelligence by (according to Wikipedia) 'a human judge engages in a natural language conversation with one human and one machine, each of which try to appear human; if the judge cannot reliably tell which is which, then the machine is said to pass the test'. While there are many researchers who refute his assessment model, there are malwares that exhibit such capabilities by deceiving their victims into believing they are humans thereby deceiving the human to surrender personal or sensitive information through "social engineering". This transgresses into another discipline of Human-Computer-Interaction (HCI) which is not covered in the present paper.

III. Use of Artificial Intelligence

A. Malware

Originally, this research started with the intent to study the characteristics of malware (eg, polymorphisms) instead of types of malware that are in existence such as worms, virus… etc. However, given the broad definition of malware, there is no universally defined standard characteristics of malware. Instead, this paper will focus on commonly known malware like email spam, virus, trojan horses and worms.

In this section, the findings on malware are organized into the following categories.

- Malware that incorporates artificial intelligence techniques such as genetic algorithms,
- Malware that have intelligent behaviours,
- Malware that have biological equivalent behaviours,
- Malware that have human like behaviours.

1) Malware with AI technologies Incorporated: There are very little literature explicitly stating malware employing artificial intelligence technologies. There is one explicitly reported virus [8] named *Zellome* that contains genetic algorithms (GA) as a form of brute-force approach to generate decryptor routine to facilitate its polymorphic behaviour. Symantec did a study into this virus and concluded its poor application of artificial intelligence technologies [9]. However its use of AI did draw some attentions.

2) Malware exhibits intelligent-like behaviours: Studies into the behaviour of malware have led researchers and anti-malware developers to note some software are exhibiting intelligence [10] such as non-predictive behaviours [11]. There are malwares like *Storm* that exhibits some forms of artificial intelligence capabilities like automatically adapting its defensive techniques to counter any measures to stop its propagation [10].

3) Malware behaving like biological equivalents: There are malwares that behave like biological equivalents / disease analogies or has attributes of artificial life. Studies [12] found that there are noticeably strong similarities between biological viruses that living organisms and their computer counterparts. For example, a study by Kienzle and Elder [11] noted that the majority of the computer worms are derivative of worms found in nature. Examples of similarities include infecting their host through an opening and replicating itself at the expense of the host. Both have abilities to spread autonomously without any human intervention. Both can be remain dormant for a period before striking. Both behaviours are becoming more malignant when combining capabilities of other like entities. An example for malware is the *Nimda* worm which is a combination of two other worms that were launched after September attack against the United States. Malware has also known to exhibit like biological parasite behaviours. Interestingly, according to Furnell and Ward [13], it has been noted that there has been a rise in malware with parasitic characteristics with less destructive payload loaded in them. The authors also noted that profit oriented motivation is the key driver in this increase. Researchers have attempted to model characteristics on the spread of malware infection using biological epidemic models. According to Chen and Ji [10], a homogeneous epidemic model was adequately modelled the propagation patterns of random-scanning worms. Some researchers have gone further to advocate that malware like viruses are possibly a form of artificial life. Artificial life have properties that include self-reproduction, information storage of its own representation, growth capabilities and evolutionary capabilities. Spafford [14] argues that computer virus exhibits close similarities to some of the defined artificial life properties like information of its self-representation. However he stops short to crediting computer virus as an artificial life as there are number of significant deficiencies found like the dependence that computer virus has on its computer host.

4) Malware behaving like humans or intelligent behaviours: There exists malware that exhibits human like behaviour. An example of such is the *IM.Myspace04.AIM* worm that managed to deceive thousands of AOL users by initiating chats with its victims using human styles of communication using shorthand phrases and slang. It lures its victim into its infectious bite by inviting them to click on a link [15]. Another example is *CyberLover* [16] found in the Russian chat forums that conducts online flirtation with intentions is to extract personal information from its victims. Typically such social engineering attacks are done by humans themselves. However *CyberLover* proves that AI malware can do likewise. This begs an answer for the question, "Could *CyberLover* possibly pass the Turing Test ?"

B. Anti-Malware

Artificial intelligence has been used extensively in anti-malware solutions to fend off malware assaults. The motivation to use artificial intelligence to empower anti-malware solutions is due to the characteristics and evolution of the intelligent malware mentioned earlier.

The survey findings of anti-malware with artificial intelligence capabilities can be grouped into the following.

- Use of artificial intelligence techniques into anti-malware solutions,
- Anti-malware solutions designed to behave like biological equivalents.

1) Anti-malware with AI techniques applied: The use of AI techniques has been largely based on the available papers or research publication. Noticeably much of the research into using AI has been focused on detection mechanisms such as Intrusion Detection Systems (IDS) or anti-malware scanners. For example, artificial neural networks [17], expert systems and fuzzy searches [18] are used to detect malware. Other forms of application of AI include identification of spam emails using natural language processors [19].

2) Anti-Malware behaving like biological equivalents: Given that malware in many instances exhibits behaviour of biological infectious equivalents, this leads to a significant amount of research into building biological equivalent defences. Capabilities like automated response and self-repair, dynamism in defences in changing attack patterns or attacker forms [20]. There is research into enhancing existing forms of anti-malware defences like Intrusion Detection System using immunological principles [20]. This area of research has also led to the study of developing a complete immune system artificially in a computer system or artificial immune systems [21] (or AIS) that attempts to detect new malware infection, analyse and remove them autonomously. The motivation to study this is that the natural immune systems since the existence of life had to deal with the imperfect world filled with harmful organisms. The natural immune system strengthens with each infectious encounter. In addition, the immune system works autonomously without any explicit intervention. This serves as an ideal model to acquire into the present day computer systems. However the research community [22] commented that purely imitating the biological immune systems may not arrive at an ideal solution as there would be specific risks associated with non-biological infection. In addition, the computing or network environment currently does not mimic closely our natural environment. However research studies gathered ([21] and [22]) also noted the differences in the objectives of information security and immune systems. Information security focuses on confidentiality, integrity, availability, accountability, and correctness with greater emphasis on confidentiality while immune system focuses on survival that is more of a combination of integrity and availability.

IV. Research Direction

Fernandez and Bureau [23] cites that the worst has yet to come as malware can further evolve technologically with the inclusion of artificial intelligence. Similar development into the use of artificial intelligence in anti-malware will likely continue in order to gain a footing over malware. Given the large community at both sides working on the advancement of malware and anti-malware, its advancement and arms race in the virtual world will continue in the foreseeable future. Wh areas of research opportunities will exist and take dominance in the use of artificial intelligence in malware and anti-malware solutions?

Future surveys of the use of artificial intelligence in malware can be quantitative with statistics. In addition intelligence assessment framework can be defined and used to assess intelligent characteristics of malware and anti-malware. For malware, specifically the ones assessed to have intelligent capabilities could be dissected further to better understand how artificial intelligence is used and publishing such findings as there are limited literature in this. Biologically inspired anti-malware solutions can be developed. A panel discussion noted that there exists a number of challenges that need to be addressed urgently [24]. One of which is the need for information security experts to have a deeper understanding on how the biological immune system functions. Also there is a need to clearly define the intention of such research direction given the objective of the information security differs from biological mechanism. Other areas yet to be considered are social engineering which incorporate HCI and psychological issues.

V. Conclusion

There is no end in sight in the war between malware and anti-malware. Both malware and anti-malware have used artificial intelligence technologies or have exhibited noticeable intelligent behaviours. The future going forward is likely to have advanced development in introducing intelligence techniques and enhanced intelligence capabilities incorporating human characteristics, knowledge and wisdom.

References

[1] (2008) TechNewsWorld website. [Online]. Available: http://www.technewsworld.com/story/61942.html?wlc=1221912220

[2] (2005) PC Magazine website. [Online]. Available: http://www.pcmag.com/article2/0,2817,1880013,00.asp

[3] F. Cohen, "Computer viruses: theory and experiments," Computers and Security, vol. 6, pp. 22-35, Feb 1987.

[4] (2007) Asia.Internet.com website. [Online]. Available: http://asia.internet.com/briefs/article.php/3678606/Estonia+Under+Russian+Cyber+Attack.htm

[5] OECD Working Party on Information Security and Privacy (WPISP) in partnership with the Asia Pacific Economic Co-operation Telecommunication and Information Working Group (APEC TEL) Security and Prosperity Steering Group (SPSG), "Malicious Software (Malware): A Security Treat to the Internet Economy," *OECD Ministerial Meeting on the Future of the Internet Economy.*, DSTI/ICCP/REG(2007)5/FINAL, Jun. 2008.

[6] V. Rao Vemuri and V. Vemuri, *Enhancing Computer Security with Smart Technology*, CRC Press, 2006.

[7] L. Steels, "The artificial life roots of artificial intelligence," Artificial Life Journal, MIT Press, Cambridge, vol 1.1, 1994.

[8] (2005) Thinkquest.org. [Online]. Available: http://library.thinkquest.org/05aug/01158/viruses.html

[9] P. Ferrie and H. Shannon, "Virus Analysis 2 – It's Zell(d)ome the one you expect," Symantec Security Response, USA, May 2005.

[10] Z. Chen and C. Ji, "Intelligent Worms: Searching for Preys," School of Electrical and Computer Engineering, Georgia Institute of Technology, USA, Jan 2006.

[11] D. Kienzle and M. Elder, "Recent Worms: A Survey and Trends," WORM '03, October 27, 2003, Washington, DC, USA.

[12] (2002) Science In Africa website. [Online]. Available: http://www.scienceinafrica.co.za/2002/october/viruses.htm

[13] S. Furnell and J. Ward, "Malware comes of age: The arrival of the true computer parasite," Network Security, vol. 2004, issue 10, pp. 11-15, Oct 2004.

[14] E. H. Spafford, "Computer Viruses as Artificial Life Journal of Artificial Life, MIT Press, 1994.

[15] (2005) PC World website. [Online]. Available: http://www.pcworld.com/article/123854/new_aim_bot_chats_spreads.html

[16] (2007) News CNET website. [Online]. Available: http://news.cnet.com/8301-13860_3-9831133-56.html

[17] J. O. Kephart et al., "Biologically Inspired Defenses Against Computer Viruses," International Joint Conference on Artificial Intelligence, vol. 14, no. 1, pp. 985-996, 1995

[18] (2007) Viruslist.com website. [Online]. Available: http://www.viruslist.com/analysis?pubid=204791972

[19] (2003) Network World website. [Online]. Available: http://www.networkworld.com/news/tech/2003/0414techupdate.html

[20] S. Forrest, S. A. Hofmeyr and A. Somayaji, "Computer Immunology," Communications of the ACM, vol. 40, no. 10, Oct 1997.

[21] G. N. Youansi, "Artificial Immune System," Communication and Operating Systems Group, Berlin University of Technology, 2006.

[22] S. Forrest, S. A. Hofmeyr and A. Somayaji, "Principles of a Computer Immune System," Proceedings of the 1997 workshop on New security paradigms, 1997.

[23] J. M. Fernandez and P. M. Bureau, "Optimising Malware," Malware'06, Phoenix, Arizona USA, April 2006.

[24] A. Somayaji, M. Locasto and J. Feyereisl, "Panel: The Future of Biologically-Inspired Security: Is There Anything Left to Learn?," New Security Paradigms Workshop 2007, North Conway, New Hampshire, USA, 2007.